A constructive proof of cut elimination for a system of full second order logic

Sandro Skansi

June 7, 2016

Abstract

In this paper we present a constructive proof of cut elimination for a system of full second order logic with the structural rules absorbed and using sets instead of sequences. The standard problem of the cutrank growth is avoided by using a new parameter for the induction, the cutweight. This technique can also be applied to first order logic.

The question of cut elimination for second and higher order logic was initially stated as the fundamental conjecture of GLC in [7], and hence became known as Takeuti's Conjecture. It was solved independently by W. Tait in 1966. [6] and D. Prawitz in 1967. [4]. Both of these proofs were semantical and did not have the form of a transformation algorithm for obtaining cutfree proofs. The problem of finding such an algorithm was stressed in [5][p. 176].

Our goal in this paper is to prove the Takeuti conjecture by constructive, proof-theoretic, means for a system of full second order logic, $\mathsf{SO}^\triangleleft$, which has the structural rules of contraction and weakening absorbed and is equivalent to Takeuti's $\mathsf{G}^1\mathsf{LC}$.

In this approach, a coverage of the relationship between first and second order terms is not needed, since no rule of the calculus presented below operates on both of them simultaneously, and therefore no problem can arise since first order and second order variable do not syntactically interact in deductions, deduction transformations, or during the procedure of cut elimination itself.

In this paper we follow the cut elimination procedure presented in [8] for first order logic and we modify this approach to be used for second order logic. Unlike [8], we consider Γ and Δ in any sequent $\Gamma \Rightarrow \Delta$ to be sets, not multisets or sequences. This approach as such is used by [3][p. 54] and [2][pp. 167-8]. Philosphically, under the intended interpretation of sequents, one could make a case that this is the true meaning of sequents

1

(as we consider e.g. $A \lor B \lor B$ to be the same as $A \lor B$), but our reason is much more pragmatic: We need this approach with sets for the proof of the closure under contraction to work. Regarding our use of sets, we will not need anything more than the axiom of extensionality in our proof and we take the version of the axiom of extensionality from [1][p. 3].

The system $\mathsf{SO}^\triangleleft$ is a standard system for full second order logic, with the following sequent rules:

$$\Gamma, F \Rightarrow F, \Delta, \ F \text{ is any formula} \qquad\qquad \Gamma, \bot \Rightarrow \Delta$$

$$\frac{\Gamma \Rightarrow \Delta, P \qquad Q, \Gamma \Rightarrow \Delta}{P \to Q, \Gamma \Rightarrow \Delta} \ \mathsf{L}\to \qquad\qquad \frac{P, \Gamma \Rightarrow \Delta, Q}{\Gamma \Rightarrow \Delta, P \to Q} \ \mathsf{R}\to$$

$$\frac{P(t), \Gamma \Rightarrow \Delta}{\forall x P(x), \Gamma \Rightarrow \Delta} \ \mathsf{L}\forall$$

$$\frac{P(T), \Gamma \Rightarrow \Delta}{\forall X P(X), \Gamma \Rightarrow \Delta} \ \mathsf{L}\forall^2 \qquad\qquad \frac{\Gamma \Rightarrow \Delta, P(a)}{\Gamma \Rightarrow \Delta, \forall x P(x)} \ \mathsf{R}\forall$$

$$\frac{\Gamma \Rightarrow \Delta, D \qquad D, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta} \ \mathsf{CUT}_{\mathsf{CS}} \qquad\qquad \frac{\Gamma \Rightarrow \Delta, P(A)}{\Gamma \Rightarrow \Delta, \forall X P(X)} \ \mathsf{R}\forall^2$$

Were t is any first order term (free or bound first order variable, function, constant), T is any second order term (second order varable (free or bound) or a formula (first or second order)), a is a free first order variable, and A is a free second order variable. In the rules $\mathsf{R}\forall$ and $\mathsf{R}\forall^2$, the respective variables a and A must not be free in the conclusion of the sequent obtained by the given rule application. In the subsequent text, we implicitly assume that the variable conditions for every $\mathsf{R}\forall$ and $\mathsf{R}\forall^2$ have been satisfied.

Definition 1 (Deduction depth). *Deduction trees are defined in the usual manner, and we follow the notation laid in [8] and write $\Pi \vdash_{\mathsf{SO}^\triangleleft} \Gamma \Rightarrow \Delta$ if and only if there is a deduction Π in $\mathsf{SO}^\triangleleft$ for the sequent $\Gamma \Rightarrow \Delta$. If Π is a deduction tree for $\Gamma \Rightarrow \Delta$, then by* deduction depth *we mean the length of the longest branch in Π. If we claim that the sequent $\Gamma \Rightarrow \Delta$ is provable by a deduction Π of depth n, we denote this by $\Pi \vdash_n \Gamma \Rightarrow \Delta$*

Definition 2 (Main formula, context, cutformula). *We define the main formula for each rule:*

- *The main formula in the axioms is the formula F.*

- *For the rule* $\mathtt{L} \to$, $P \to Q$ *is the main formula. Similarly for* $\mathtt{R} \to$, $\mathtt{L}\forall$, $\mathtt{L}\forall^2$, $\mathtt{R}\forall^2$. $\mathtt{R}\forall$.

- *The cut rule has no main formula, but it has the cut formula, viz. D.*

All other formulas in the sequents are called the context of the given rule.

Definition 3. *We define the other connectives in the usual manner:* $\bot := \forall XX$, $\neg P := P \to \bot$, $P \vee Q := \neg P \to Q$, $P \wedge Q := \neg(\neg P \vee \neg Q)$, $\exists x P(x) := \neg \forall x \neg P(x)$, $\exists X P(X) = \neg \forall X \neg P(X)$

Definition 4 (Weight, cutweight)**.** *Let* Π *be a deduction, and let* $\kappa_0, \kappa_1, \ldots, \kappa_m$ *denote the nodes (i.e. sequents) in the deduction tree, and let* $\mathfrak{A}$ *denote the set of all leaves. We define the function* $\pounds$ *which we call the weight of the node (sequent) such that: the weight of a leaf is 1. If from a node (sequent) of weight i we deduce the next sequent by a single premise rule, the weight of the sequent is* $(i + i)$*. If from two nodes (sequents) of the weight i and j we deduce the next sequent by a double premiss rule, the weight of the sequent is* $(i + j)$*.*

$$\pounds(\kappa_k) := \begin{cases} 1, \kappa_k \in \mathfrak{A} \\ i + i, \kappa \notin \mathfrak{A}, \pounds(\kappa_{k-1}) = i \\ i + j, \kappa \notin \mathfrak{A}, \pounds(\kappa'_{k-1}) = i, \pounds(\kappa''_{k-1}) = j \end{cases}$$

We define the weight of a rule occurrence in the deduction tree as the weight of the conclusion of that rule. The weight of an instance of the cut rule is the weight of the conclusion of that instance of the cut rule. The weight of the whole deduction tree is the weight of its conclusion and is denoted by $\pounds(\Pi)$*. The* cutweight *of a deduction tree* $\pounds_{\mathtt{CUT}}(\Pi)$ *is the weight of the cut with the maximal weight in the deduction tree. If there are no cuts in the deduction tree we set* $\pounds_{\mathtt{CUT}}(\Pi) = 0$*.*

We give an example of a deduction tree with weights:

$$\cfrac{\cfrac{1}{(1+1) = 2}}{\cfrac{((1+1) + (1+1)) = 4}{\quad}} \quad \cfrac{\cfrac{1}{(1+1) = 2} \quad \cfrac{\cfrac{1}{(1+1) = 2}}{((1+1) + (1+1)) = 4}}{((1+1) + ((1+1) + (1+1))) = 6}$$
$$(((1+1) + (1+1)) + ((1+1) + ((1+1) + (1+1)))) = 10$$

We note that the maximum weight for deduction Π where the depth $d(\Pi) = n$ is $2^{(d(\Pi)-1)}$. If a deduction Π with depth $d(\Pi) = n$ is transformed into a deduction Π' with depth $d(\Pi') = n$ but with more branches and all the lengths of the branches are equal to the length of the longest branch (which is by the above definition the depth of the deduction), the weight of the deduction is unchanged, whereas whenever a deduction under a given transformation branches with unequal branches, the weight of the resulting deduction of the same depth as the original is decreased.

As we have defined the weight and related concepts, we prove below three lemmas, which will enable us to use induction on depth in some situations, and know that the the fact that the depth ha decrased ensures the weight has also decreased.

Definition 5 (Subdeduction). *Let Π be a deduction with the endsequent (conclusion) $\Gamma \Rightarrow \Delta$ ending with a one-premise rule, using as the premise $\Gamma' \Rightarrow \Delta'$. Then we call the deduction of that premise $\Pi' \vdash \Gamma \Rightarrow \Delta$ the immediate subdeduction of Π. We define the case when two premises are present in the same way, only then both premise deductions Π' and Π'' are immediate subdeductions of Π.*

If there more than one rule needed to derive $\Gamma \Rightarrow \Delta$ from $\Gamma' \Rightarrow \Delta'$, then we call Π' just a subdeduction of Π. We expand this definition to two premise rules in a natural way.

Lemma 1. *If Π' is a subdeduction or immediate subdeduction of Π, then $\pounds(\Pi') \leq \pounds(\Pi)$.*

Proof. Obvious from the definition of weight. □

Lemma 2. *Let Π and Π' be deductions. If Π ends with a one premise rule and $d(\Pi) \geq d(\Pi')$ or both Π and Π' end with a two premise rule and $d(\Pi) \geq d(\Pi')$, then $\pounds(\Pi) \geq \pounds(\Pi')$.*

Proof. By induction on the depth of Π.

Basis: If $d(\Pi) = 1$, then Π is an axiom. As $d(\Pi') \leq d(\Pi)$, it follows that $d(\Pi') \leq 1$, so Π' is an axiom. According to the definition of weight, $\pounds(\Pi') = 1$.

Induction step: Let the statement hold for deductions of depth n (we denote them by Π_n and Π'_n). We prove that the statement holds for deductions of depth $n + 1$ (we denote them by Π_{n+1} and Π'_{n+1}).
We distinguish tree cases:

1. Π_{n+1} is obtained by applying a one premise rule to the conclusion of Π_n and Π'_{n+1} is obtained applying a one premise rule to the conclusion Π'_n. Then, according to the definition of weight $\pounds(\Pi_{n+1}) = 2 \cdot \pounds(\Pi_n)$ and $\pounds(\Pi'_{n+1}) = 2 \cdot \pounds(\Pi'_n)$. According to the induction hypothesis $\pounds(\Pi'_n) \leq \pounds(\Pi_n)$, and it follows that $\pounds(\Pi'_{n+1}) \leq \pounds(\Pi_{n+1})$.

2. Π_{n+1} is obtained by applying a one premise rule to the conclusion of Π_n and Π'_{n+1} is obtained with a two premise rule with one of the premises being the conclusion of Π'_n, while the deduction of the second premise is denoted by Σ. W.l.o.g. we can assume that Π'_n is the left premise and $d(\Sigma) \leq d(\Pi'_n)$. By the weight definition it follows that $\pounds(\Pi_{n+1}) = 2 \cdot \pounds(\Pi_n)$ and $\pounds(\Pi'_{n+1}) = \pounds(\Pi'_n) + \pounds(\Sigma)$. By the induction hypothesis $\pounds(\Pi'_n) \leq \pounds(\Pi_n)$. As $d(\Sigma) \leq d(\Pi'_n)$, by applying the IH we get $\pounds(\Sigma) \leq \pounds(\Pi'_n)$. From this follows: $2 \cdot \pounds(\Pi_n) \geq \underbrace{\pounds(\Pi'_n)}_{\leq \pounds(\Pi_n)} + \underbrace{\pounds(\Sigma)}_{\leq \pounds(\Pi'_n) \leq \pounds(\Pi_n)}$.

   Therefore, $\pounds(\Pi'_{n+1}) \leq \pounds(\Pi_{n+1})$.

3. Let Π_{n+1} and Π'_{n+1} end with two premise rules and let Σ and Σ' denote the deductions of the premises with lower depth.

   The deductions of the premises with greater depth are denoted by Π_n and Π'_n while the deductions of the premises with lesser depth are denoted by Σ i Σ'. If they are of equal depth, then Π_n and Π'_n denote the left premises.

   From the assumption of the lemma follows that $d(\Sigma) \leq d(\Pi_n)$ i $d(\Sigma') \leq d(\Pi'_n)$. By the induction hypothesis follows: $\pounds(\Pi'_n) \leq \pounds(\Pi_n)$. From the lemma assumption follows that $d(\Sigma) \leq d(\Sigma')$ by applying the IH follows that $\pounds(\Sigma) \geq \pounds(\Sigma')$. As $\pounds(\Pi_{n+1}) = \pounds(\Pi_n) + \pounds(\Sigma)$ and $\pounds(\Pi'_{n+1}) = \pounds(\Pi'_n) + \pounds(\Sigma')$ follows $\pounds(\Pi_{n+1}) \geq \pounds(\Pi_{n+1})$.

$\square$

    The next set of lemmas are inteded to ensure we have closure under substitution, contraction and weakening in $\mathsf{SO}^{\lhd}$. These lemmas serve two purposes: (1) We use them in the proof of cut elimination itself, so they have technical importance, and (2) when we have the desired closure properties, we have the proof of the admissibility of such rules in $\mathsf{SO}^{\lhd}$, which shows that the system $\mathsf{SO}^{\lhd}$ with the structural rules absorbed is deductively equivalent to a system which does not have the structural rules absorbed, but rather lists them among all the other rules (as e.g. Takeuti's $\mathsf{G}^1\mathsf{LC}$).

Lemma 3 (Closure for $SO^\lhd$ under weakening). *$SO^\lhd$ is closed under weakening without the increase of depth, i.e. if $\vdash_n \Gamma \Rightarrow \Delta$, then $\vdash_n \Gamma' \Rightarrow \Delta'$, where $\Gamma \subseteq \Gamma'$ and $\Delta \subseteq \Delta'$.*

Proof. Induction on the deduction depth. Axioms:

If $\Gamma \Rightarrow \Delta$ is an axiom, then it has the form $\Gamma'', F \Rightarrow F, \Delta''$ or $\Gamma'', \bot \Rightarrow \Delta$. We consider only the first case, the second is treated in the same manner. Since $\Gamma \subseteq \Gamma'$ and $\Delta \subseteq \Delta'$, it follows $F \in \Gamma'$, $F \in \Delta'$. Then $\Gamma' \Rightarrow \Delta'$ is also an axiom.

Induction step, by cases.

1. Let $\Gamma' \Rightarrow \Delta', P$ and $Q, \Gamma' \Rightarrow \Delta'$ be of depth n . Then by application of $\mathtt{L} \to$ we get $P \to Q, \Gamma' \Rightarrow \Delta'$, of depth $n + 1$.

2. Let $\Gamma', P \Rightarrow \Delta', Q$ be of depth n. Then by application of $\mathtt{R} \to$ we get $\Gamma' \Rightarrow \Delta', P \to Q$, of depth $n + 1$.

3. Let $\Gamma', P(t) \Rightarrow \Delta'$ be of depth n. Then by application of $\mathtt{L}\forall$ we get $\Gamma', \forall x P \Rightarrow \Delta'$, of depth $n + 1$.

4. Let $\Gamma' \Rightarrow \Delta', P(a)$ be of depth n. Then by application of $\mathtt{R}\forall$ we get $\Gamma' \Rightarrow \Delta', \forall x P$, of depth $n + 1$.

5. Let $\Gamma', P(T) \Rightarrow \Delta'$ be of depth n. Then by application of $\mathtt{L}\forall^2$ we get $\Gamma', \forall X P \Rightarrow \Delta'$, of depth $n + 1$.

6. Let $\Gamma' \Rightarrow \Delta', P(A)$ be of depth n. Then by application of $\mathtt{R}\forall^2$ we get $\Gamma' \Rightarrow \Delta', \forall X P$, of depth $n + 1$.

As depth has not increased, the weight has also not increased by lemma 2 (for cases 2-6) and lemma 1 (for case 1). $\qquad\square$

Lemma 4 (First order term substitution). *If $\Pi \vdash \Gamma \Rightarrow \Delta$, then $\Pi[x/t] \vdash \Gamma[x/t] \Rightarrow \Delta[x/t]$, where t is free for x in $\Gamma \Rightarrow \Delta$ and does not contain free variables for any occurence of the rule $\mathtt{R}\forall$ in the deduction Π. This substitution preserves the depth and weight of the deduction.*

Proof. By induction on the depth of deductions.
Basis, deductions of depth 1. We have two cases:

- $\Gamma \Rightarrow \Delta$ is an axiom. Then $\Gamma \Rightarrow \Delta$ is of the form $\Gamma', F \Rightarrow F, \Delta'$, where F is an arbitrary formula. We distinguish two subcases: (i) x occurs

in F, so F is of the form $F(x)$. The result of the substitution is then $F(t)$, which gives us $\Gamma', F(t) \Rightarrow F(t), \Delta'$, which is an instance of an axiom. (ii) x occurs in the context. As F remains unchanged, the sequent is still an instance of an axiom.

- $\Gamma \Rightarrow \Delta$ is an instance of the rule for $\bot$. The proof is the same as for the case above, with slight modifications.

Induction step. We assume the claim holds for deductions of depth n and prove for deductions of depth $n + 1$. We give a proof by cases, restricting our attention to proofs which end up with the rules L$\forall$ and R$\forall$.

(1) L$\forall$. By the induction hypothesis the substitution has already been carried out in the premise $P(s)[x/t], \Gamma[x/t] \Rightarrow, \Delta[x/t]$. As t has to be free for instances of the rule L$\forall$, and $t \notin FV(s)$, we apply L$\forall$ to the premise and get $\forall y P[x/t], \Gamma[x/t] \Rightarrow, \Delta[x/t]$.

(2) R$\forall$. Symmetrical to the case (1) above. $\qquad\qquad\square$

Lemma 5 (Second order term substitution)**.** *Let T and S be any second order terms. If $\Pi \vdash \Gamma \Rightarrow \Delta$, then $\Pi[S/T] \vdash \Gamma[S/T] \Rightarrow \Delta[S/T]$, where T is free for S in $\Gamma \Rightarrow \Delta$ and does not contain free variables i which are used as free variables in any instance of the rule R$\forall^2$ in the deduction Π and S is not equal to the free variable which is used in any instance of the rule R$\forall^2$ in Π. This substitution preserves the depth and the weight of the deduction, but not the length of the formulas in the conclusion $\Gamma \Rightarrow \Delta$.*

Proof. Induction on the deduction depth.

If $\Gamma \Rightarrow \Delta$ is an axiom, then $\Gamma \Rightarrow \Delta$ is of the form $\Gamma', P \Rightarrow P, \Delta'$, where P is an arbitrary formula. We distinguish two subcases: (i) S occurs in P, so P is of the form $P(S)$. The result of the substitution is then $P(T)$, which gives $\Gamma', P(T) \Rightarrow P(T), \Delta'$. The result is again an axiom. (ii) S occurs in the context. Then the formula P remains unchanged, and we have an axiom.

Induction step. We have the following cases:

1. Case L $\rightarrow$. By the induction hypothesis the substitution is already carried out in the premisses $\Gamma[S/T] \Rightarrow \Delta[S/T]P[S/T]$ and $\Gamma[S/T], Q[S/T] \Rightarrow \Delta[S/T]]$. By L $\rightarrow$ we have $\Gamma[S/T], P[S/T] \rightarrow Q[S/T] \Rightarrow \Delta[S/T]$.

2. Case R $\rightarrow$. By the induction hypothesis the substitution is already carried out in the premiss $\Gamma[S/T], P[S/T] \Rightarrow \Delta[S/T]Q[S/T]$. By R $\rightarrow$ we have $\Gamma[S/T] \Rightarrow \Delta[S/T], P[S/T] \rightarrow Q[S/T]$.

3. Case L∀. By the induction hypothesis the substitution is already carried out in the premiss $P(t)[S/T], \Gamma[S/T] \Rightarrow \Delta[S/T]$. As t is an arbitrary first order term, we apply L∀ and get $\forall x P(x)[S/T], \Gamma[S/T] \Rightarrow \Delta[S/T]$

4. Case R∀. By the induction hypothesis the substitution is already carried out in the premiss $\Gamma[S/T] \Rightarrow \Delta[S/T], P(a)[S/T]$. We are free to presume that $a \notin FV(T)$, apply R∀ and get $\Gamma[S/T] \Rightarrow \Delta[S/T], \forall x P(x)[S/T]$

5. Case L∀². By the induction hypothesis the substitution is already carried out in the premiss $P(S)[S/T], \Gamma[S/T] \Rightarrow \Delta[S/T]$, where S is an arbitrary second order term. We apply L∀² to the premiss and get $\forall Y P(Y)[S/T], \Gamma[S/T] \Rightarrow, \Delta[S/T]$.

6. Case R∀². By the induction hypothesis the substitution is already carried out in the premiss $\Gamma[S/T] \Rightarrow \Delta[S/T], P(A)[S/T]$, where A is a free second order variable. We are free to assume $A \notin FV(T)$ and $A \neq T$ so we apply R∀² and get $\Gamma[S/T] \Rightarrow \Delta[S/T], \forall X P(X)[S/T]$. We note that A may be equal to S.

As depth has not increased, the weight has also not increased by lemma 2 (cases 2-6) and lemma 1 (case 1).

□

As noted earlier, there is no need to address the interaction between first and second order terms, since they do not syntactically interact. What is meant by this is that in any given rule application, either a change is made on a first order term, or a second order term, but not both simultaneously. As such, in our approach, first order and second order terms do not really interact, and therefore we don not address this interaction.

Lemma 6 (Closure under contraction). *Let $\vdash$ denote deducibility in* SO◁. *The following holds: (i) If $\vdash_k P, P, \Gamma \Rightarrow \Delta$, then $\vdash_k P, \Gamma \Rightarrow \Delta$. (ii) if $\vdash_k \Gamma \Rightarrow \Delta, P, P$, then $\vdash_k \Gamma \Rightarrow \Delta, P$. Closure under contraction preserves the weight of the deduction.*

Proof. We address only case (i). Let $\Gamma' = \Gamma \cup \{P\} \cup \{P\}$ and let $\Gamma'' = \Gamma \cup \{P\}$. By the axiom of extensionality $\Gamma' = \Gamma''$. As extensionality is not a formal rule and does not use up a deduction step, no formal change is made to the deduction, so both the depth and weight are trivially preserved.

□

The previous lemma is the only place where we use the fact that sequents have sets and not multisets or sequences. As noted earlier we are not alone in adopting this approach, but this is needed to prove this lemma. One could hope for a proof using multisets or sequences, but this argument is immaterial, since we do not actually need multisets nor sequences. [1]

Theorem 1. *Cut elimination holds for* $\mathsf{SO}^\triangleleft$.

Proof. We prove by induction on $\mathcal{L}_{\mathtt{CUT}}(\Pi)$. We have the following cases:

1. At least one of the Π_0, Π_1 is an axiom.

2. Π_0 and Π_1 are not axioms, and in at least one premiss deduction tree the cutformula is not the main formula of the previous rule.

3. The cutformula is the main formula on both sides.

Case 1.
Subcase 1.*a*. Π_0 is an axiom and D is not the main formula in Π_0. Then Π is of the form

$$\frac{P,\Gamma \Rightarrow \Delta, P, D \qquad \overset{\textstyle\Pi_1}{D, (P), \Gamma \Rightarrow \Delta, (P)}}{P, \Gamma \Rightarrow \Delta, P} \ \mathtt{CUT_{CS}}$$

As the conclusion is an axiom (P is present on both sides), for Π^* we use the conclusion which is an axiom.[2] The cutweight is reduced to 0 because the cut is eliminated.

Subcase 1.*b*. Π_0 is an axiom and the main formula in Π_0 is the cutformula (P).

$$\frac{P,\Gamma \Rightarrow \Delta, P \qquad \overset{\textstyle\Pi_1}{P, (P), \Gamma \Rightarrow \Delta}}{P, \Gamma \Rightarrow \Delta} \ \mathtt{CUT_{CS}}$$

---

[1] If the need for a multiset arises, we can rename the additional copy, e.g. to simulate the multiset $\{A, B, B\}$ we use the set $\{A, B, B_0\}$. This is possible in every case except where we subsequently need to apply contraction, but this case is nonexistent. As an ordered n-tuple can be defined using sets, the case when a sequence might be needed is also covered.

[2] A few notes. In the left premiss Π_0, D is part of the context but we isolate it since it is the cut formula. At the conclusion Π_1, on the right hand side, (P) is part of the context, since here CUT_{CS} is used so the contexts must be the same on both sides.

9

We get Π^* by applying closure under contraction to Π_1. The cutweight is reduced according to lemma 1.

Subcase 1.c. Π_1 is an axiom and D is not the main formula in Π_1. Then Π is of the form

$$
\dfrac{\begin{array}{c}\Pi_0\\ (P),\Gamma \Rightarrow \Delta,(P),D \end{array} \qquad D,P,\Gamma \Rightarrow \Delta,P}{P,\Gamma \Rightarrow \Delta,P} \; CUT_{CS}
$$

As the conclusion is an axiom (P is the main formula on both sides), for Π^* we use the conclusion which is an axiom. The cutweight is reduced to 0 because the cut is eliminated.

Subacase 1.d. Π_1 is an axiom an the main formula in Π_1 is the cutformula P.

$$
\dfrac{\begin{array}{c}\Pi_0\\ \Gamma \Rightarrow \Delta,(P),P \end{array} \qquad P,\Gamma \Rightarrow \Delta,P}{\Gamma \Rightarrow \Delta,P} \; \text{CUT}_{\text{CS}}
$$

Π^* is obtained by applying closure under contraction to Π_1. The cutweight is reduced according to lemma 1.

Case 2.

Π_0 and Π_1 are not axioms, and the cut formula is not principal and the cutformula is not principal on at least one side. W.L.O.G. we can assume the cutformula is not principal on either side

Subcase 2.a. The deduction of both premises end with a one premise rule. Then Π has the form

$$
\dfrac{\dfrac{\begin{array}{c}\Pi_{00}\\ \Gamma' \Rightarrow \Delta',D \end{array}}{\Gamma \Rightarrow \Delta,D}\text{RULE}' \qquad \dfrac{\begin{array}{c}\Pi_{10}\\ \Gamma'',D \Rightarrow \Delta'' \end{array}}{\Gamma,D \Rightarrow \Delta}\text{RULE}''}{\Gamma \Rightarrow \Delta} \; \text{CUT}_{\text{CS}}
$$

We transform the deduction into

$$
\dfrac{\dfrac{\dfrac{\begin{array}{cc}\Pi_{00}[\Gamma'' \Rightarrow \Delta''] & \Pi_{10}[\Gamma', \Rightarrow \Delta']\\ \Gamma',\Gamma'' \Rightarrow \Delta',\Delta''D & \Gamma',\Gamma'',D \Rightarrow \Delta',\Delta''\end{array}}{\Gamma',\Gamma'' \Rightarrow \Delta',\Delta''}\text{CUT}_{\text{CS}}}{\Gamma,\Gamma'' \Rightarrow \Delta,\Delta''}\text{RULE}'}{\Gamma,\Gamma \Rightarrow \Delta,\Delta}\text{RULE}''
$$

We apply closure under contraction and get the desired deduction. We repeat the transformation and note the weights of the formulas. The starting deduction was:

$$
\cfrac{
\cfrac{
\begin{array}{c} \Pi_{00} \\ \mathcal{L}(\Gamma' \Rightarrow \Delta', D) = i \end{array}
}{\mathcal{L}(\Gamma \Rightarrow \Delta, D) = i + i}\ \text{\scriptsize RULE}'
\qquad
\cfrac{
\begin{array}{c} \Pi_{10} \\ \mathcal{L}(\Gamma'', D \Rightarrow \Delta'') = j \end{array}
}{\mathcal{L}(\Gamma, D \Rightarrow \Delta) = j + j}\ \text{\scriptsize RULE}''
}{\mathcal{L}(\Gamma \Rightarrow \Delta) = 2i + 2j}\ \text{\scriptsize CUT}_{\text{CS}}
$$

Which was transformed into the following deduction, but we note that we only follow the weight of the conclusion of the cut, since this is the parameter for the induction. As we have shown, weakening and contraction do not increase the weight of any sequent.

$$
\cfrac{
\cfrac{
\begin{array}{c} \Pi_{00}[\Gamma'' \Rightarrow \Delta''] \\ \mathcal{L}(\Gamma', \Gamma'' \Rightarrow \Delta', \Delta''D) = i \end{array}
\qquad
\begin{array}{c} \Pi_{10}[\Gamma', \Rightarrow \Delta'] \\ \mathcal{L}(\Gamma', \Gamma'', D \Rightarrow \Delta', \Delta'') = j \end{array}
}{\mathcal{L}(\Gamma', \Gamma'' \Rightarrow \Delta', \Delta'') = i + j}\ \text{\scriptsize CUT}_{\text{CS}}
}{
\cfrac{* * *}{* * *}\ \text{\scriptsize RULE}''
}\ \text{\scriptsize RULE}'
$$

As obviously $(i + j) < (2i + 2j)$, the cutweight has been reduced and we apply the induction hypothesis. The case where RULE' and RULE'' have two premisses is treated in the same manner.

Subcase 2.b. The deduction of one premise ends with a two premise rule. W.L.O.G. we can assume that Π_0 end with a two premiss rule RULE

$$
\cfrac{
\cfrac{
\begin{array}{c} \Pi_{00} \\ \Gamma' \Rightarrow \Delta', D : \mathcal{L} = i \end{array}
\qquad
\begin{array}{c} \Pi_{01} \\ \Gamma'' \Rightarrow \Delta'', D : \mathcal{L} = j \end{array}
}{\Gamma \Rightarrow \Delta, D : \mathcal{L} = i + j}\ \text{\scriptsize RULE}
\qquad
\begin{array}{c} \Pi_1 \\ D, \Gamma \Rightarrow \Delta : \mathcal{L} = k \end{array}
}{\Gamma \Rightarrow \Delta : \mathcal{L} = (i + j) + k}\ \text{\scriptsize CUT}_{\text{CS}}
$$

We permute the cut over RULE and get Π'

$$
\cfrac{
\cfrac{
\begin{array}{c} \Pi_{00}[\Gamma \Rightarrow \Delta] \\ \Gamma\Gamma' \Rightarrow \Delta\Delta'D : \mathcal{L} = i \end{array}
\quad
\begin{array}{c} \Pi_1[\Gamma' \Rightarrow \Delta'] \\ D\Gamma\Gamma' \Rightarrow \Delta\Delta' : \mathcal{L} = k \end{array}
}{\Gamma\Gamma' \Rightarrow \Delta\Delta' : \mathcal{L} = i + k}
\qquad
\cfrac{
\begin{array}{c} \Pi_{01}[\Gamma \Rightarrow \Delta] \\ \Gamma\Gamma'' \Rightarrow \Delta\Delta''D : \mathcal{L} = j \end{array}
\quad
\begin{array}{c} \Pi_1[\Gamma'' \Rightarrow \Delta''] \\ D\Gamma\Gamma'' \Rightarrow \Delta\Delta'' : \mathcal{L} = k \end{array}
}{\Gamma\Gamma'' \Rightarrow \Delta\Delta'' : \mathcal{L} = j + k}\ \text{\scriptsize CUT}_{\text{CS}}
}{\Gamma\Gamma \Rightarrow \Delta\Delta}\ \text{\scriptsize RULE}
$$

We apply closure under contraction and get the deduction of the same sequent. In the transformed deduction we have two cuts, with weights $i + k$ and $j + k$, and thus the $\mathcal{L}_{\texttt{CUT}}(\Pi') = max((i+k), (j+k))$, while in the original deduction the cutweight was $i + j + k$. Since $max((i+k), (j+k)) < i+j+k$, the cutweight has been reduced and we apply the induction hypothesis.

Subcase 2.c. The deductions of both premises end with a two premise rule. Since the only two premise rules we have are $\texttt{L} \rightarrow$ and $\texttt{CUT}_{\texttt{CS}}$, and we assumed all previous cuts have been eliminated, the rules must be $\texttt{L} \rightarrow$. Therefore we transform the deduction Π

$$\cfrac{\overset{\Pi_A}{\Gamma, R \rightarrow S, P \rightarrow Q \Rightarrow \Delta, D : \mathcal{L} = i + j} \qquad \overset{\Pi_B}{\Gamma, P \rightarrow Q, R \rightarrow S, D \Rightarrow \Delta : \mathcal{L} = k + l}}{\Gamma, P \rightarrow Q, R \rightarrow S \Rightarrow \Delta : \mathcal{L} = i + j + k + l} \; \texttt{CUT}_{\texttt{CS}}$$

Where Π_A is:

$$\cfrac{\overset{\Pi_{00}}{\Gamma, R \rightarrow S \Rightarrow \Delta, P, D : \mathcal{L} = i} \qquad \overset{\Pi_{01}}{\Gamma, R \rightarrow S, Q \Rightarrow \Delta, D : \mathcal{L} = j}}{\Gamma, R \rightarrow S, P \rightarrow Q \Rightarrow \Delta, D : \mathcal{L} = i + j} \; \texttt{L} \rightarrow$$

Where Π_B is:

$$\cfrac{\overset{\Pi_{10}}{\Gamma, P \rightarrow Q, D \Rightarrow \Delta, R : \mathcal{L} = k} \qquad \overset{\Pi_{11}}{\Gamma, P \rightarrow Q, D, S \Rightarrow \Delta : \mathcal{L} = l}}{\Gamma, P \rightarrow Q, R \rightarrow S, D \Rightarrow \Delta : \mathcal{L} = k + l} \; \texttt{L} \rightarrow$$

We transform Π and get Π':

$$\cfrac{\overset{\Pi'_A}{\Gamma, R \rightarrow S, P \rightarrow Q \Rightarrow \Delta, P : \mathcal{L} = i + k + l} \qquad \overset{\Pi'_B}{\Gamma, P \rightarrow Q, R \rightarrow S, Q \Rightarrow \Delta : \mathcal{L} = j + k + l}}{\Gamma, R \rightarrow S, P \rightarrow Q \Rightarrow \Delta, D} \; \texttt{L} \rightarrow$$

Where Π'_A is:

$$\cfrac{\overset{\Pi_{00}[P \rightarrow Q \Rightarrow]}{\Gamma, P \rightarrow Q, R \rightarrow S \Rightarrow \Delta, P, D : \mathcal{L} = i} \qquad \cfrac{\overset{\Pi_{10}[\Rightarrow P]}{\Gamma, P \rightarrow Q, D \Rightarrow \Delta, R, P : \mathcal{L} = k} \qquad \overset{\Pi_{11}[\Rightarrow P]}{\Gamma, P \rightarrow Q, D, S \Rightarrow \Delta, P : \mathcal{L} = l}}{\Gamma, P \rightarrow Q, R \rightarrow S, D \Rightarrow \Delta, P : \mathcal{L} = k + l} \; \texttt{L} \rightarrow}{\Gamma, R \rightarrow S, P \rightarrow Q \Rightarrow \Delta, P : \mathcal{L} = i + k + l} \; \texttt{CUT}_{\texttt{CS}}$$

And Π'_B is

$$\cfrac{\cfrac{\Pi_{01}[P \to Q \Rightarrow]}{\Gamma, P \to Q, R \to S, Q \Rightarrow \Delta, D : \pounds = j} \qquad \cfrac{\cfrac{\Pi_{10}[Q \Rightarrow]}{\Gamma, Q, P \to Q, D \Rightarrow \Delta, R : \pounds = k} \qquad \cfrac{\Pi_{11}[Q \Rightarrow]}{\Gamma, Q, P \to Q, D, S \Rightarrow \Delta : \pounds = l}}{\Gamma, Q, P \to Q, R \to S, D \Rightarrow \Delta : \pounds = k+l} \text{ L} \to}{\Gamma, Q, P \to Q, R \to S \Rightarrow \Delta : \pounds = j+k+l} \text{ CUT}_{\text{CS}}$$

With this transformation we get from $\pounds_{\text{CUT}}(\Pi) = i + j + k + l$ the cutweight for Π' which is $\pounds_{\text{CUT}}(\Pi') = max(i + k + l, j + k + l)$, and clearly $\pounds_{\text{CUT}}(\Pi') < \pounds_{\text{CUT}}(\Pi)$.

Note: RULE cannot be an instance of cut CUT$_{\text{CS}}$, since we assumed that all previous cuts have been eliminated.

Case 3. The cutformula is the main formula in Π_0 and Π_1

Subcase 3.a. $D \equiv D_0 \to D_1$.

The deduction Π

$$\cfrac{\cfrac{\cfrac{\Pi_{00}}{D_0, \Gamma \Rightarrow \Delta, D_1 : \pounds = i}}{\Gamma \Rightarrow \Delta, D_0 \to D_1 : \pounds = 2i} \text{ R} \to \qquad \cfrac{\cfrac{\Pi_{10}}{\Gamma \Rightarrow \Delta, D_0 : \pounds = j} \qquad \cfrac{\Pi_{11}}{\Gamma, D_1 \Rightarrow \Delta : \pounds = k}}{\Gamma, D_0 \to D_1 \Rightarrow \Delta : \pounds = j+k} \text{ L} \to}{\Gamma \Rightarrow \Delta : \pounds = 2i + j + k} \text{ CUT}_{\text{CS}}$$

Is transformed into Π'

$$\cfrac{\cfrac{\Pi_{10}}{\Gamma \Rightarrow \Delta, D_0 : \pounds = j} \qquad \cfrac{\cfrac{\Pi_{00}}{\Gamma, D_0 \Rightarrow \Delta, D_1 : \pounds = i} \qquad \cfrac{\Pi_{11}[D_0 \Rightarrow]}{\Gamma, D_0, D_1 \Rightarrow \Delta : \pounds = k}}{\Gamma, D_0 \Rightarrow \Delta : \pounds = i + k} \text{ CUT}_{\text{CS}}}{\Gamma \Rightarrow \Delta : \pounds = j + i + k} \text{ CUT}_{\text{CS}}$$

We note that we have two cuts, but the cutweight has decreased, so we apply the induction hypothesis.

Subcase 3.b. $D \equiv \forall x D_0(x)$. The deduction Π $(a \notin FV(D_0, \Gamma, \Delta, t))$

$$\cfrac{\cfrac{\cfrac{\Pi_{00}}{\Gamma \Rightarrow \Delta, D_0(a) : \pounds = i}}{\Gamma \Rightarrow \Delta, \forall x D_0(x) : \pounds = 2i} \text{ R}\forall \qquad \cfrac{\cfrac{\Pi_{10}}{\Gamma, D_0(t) \Rightarrow \Delta : \pounds = j}}{\Gamma, \forall x D_0(x) \Rightarrow \Delta : \pounds = 2j} \text{ L}\forall}{\Gamma \Rightarrow \Delta : \pounds = 2i + 2j} \text{ CUT}_{\text{CS}}$$

is transformed into Π'

13

$$\frac{\begin{array}{cc} \Pi_{00}[a/t] & \Pi_{10} \\ \Gamma \Rightarrow \Delta, D_0(t) : \mathcal{L} = i \quad & \Gamma, D_0(t) \Rightarrow \Delta : \mathcal{L} = j \end{array}}{\Gamma \Rightarrow \Delta : \mathcal{L} = i+j} \text{ CUT}_{\text{CS}}$$

The cutweight has decreased so we eliminate the cut with the induction hypothesis.

Subacase 3.c. $D \equiv \forall X D_0(X)$. The deduction Π $(A \notin FV(D_0, \Gamma, \Delta, T))$

$$\frac{\dfrac{\begin{array}{c}\Pi_{00}\\ \Gamma \Rightarrow \Delta, D_0(A) : \mathcal{L} = i\end{array}}{\Gamma \Rightarrow \Delta, \forall X D_0(X) : \mathcal{L} = 2i} \text{R}\forall \quad \dfrac{\begin{array}{c}\Pi_{10}\\ \Gamma, D_0(T) \Rightarrow \Delta : \mathcal{L} = j\end{array}}{\Gamma, \forall X D_0(X) \Rightarrow \Delta : \mathcal{L} = 2j} \text{L}\forall}{\Gamma \Rightarrow \Delta : \mathcal{L} = 2i + 2j} \text{ CUT}_{\text{CS}}$$

is transformed into Π'

$$\frac{\begin{array}{cc} \Pi_{00}[A/T] & \Pi_{10} \\ \Gamma \Rightarrow \Delta, D_0(T) : \mathcal{L} = i \quad & \Gamma, D_0(T) \Rightarrow \Delta : \mathcal{L} = j \end{array}}{\Gamma \Rightarrow \Delta : \mathcal{L} = i+j} \text{ CUT}_{\text{CS}}$$

The cutweight has decreased so we eliminate the cut with the induction hypothesis

$\square$

In this paper we have given a constructive proof of cut elimination for a system of second order logic with the structural rules absorbed. We have done this by using a new parameter for induction, viz the cutweight. By doing so we have avoided the problems arising from the impredicativity of second order logic. It is our belief that this procedure can be applied to all finite order logic since the problems are not essentially different from the problem arising from impredicativity in second order logic.

References

[1] T. Jech. *Set Theory.* Springer, New York, 2002.

[2] G. S. Boolos & J. P. Burgess & R. C. Jeffrey. *Computability and Logic.* Cambridge University Press, New York, 2007.

[3] W. Pohlers. *Proof Theory: An Introduction.* Springer, New York, 2002.

[4] D. Prawitz. Completeness and hauptsatz for second order logic. *Theoria*, 33:246–258, 1967.

[5] A. Pillay S. R. Buss, A. S. Kechris and R. Shore. The prospects for mathematical logic in the twenty-first century. *The Bulletin of Symbolic Logic*, 7(2):169–196, 2001.

[6] W. W. Tait. A nonconstructive proof of Gentzen's Haupsatz for second order predicate logic. *Bulletin of the American Mathematical Society*, 72(6):980–83, 1966.

[7] G. Takeuti. On a generalized logic calculus. *Japanese Journal of Mathematics*, 23:39–96, 1953.

[8] A. S. Troelstra and H. Schwichtenberg. *Basic Proof Theory*. Cambridge University Press, Cambridge, 2000.